DOI: 10.1111/fire.12406

ORIGINAL ARTICLE

The Financial Review WILEY

The price impact of tweets: A high-frequency study

Ni Yang[1] | Adrian Fernandez-Perez[3] | Ivan Indriawan[2]

[1]Economics and Finance Department, Auckland University of Technology, Auckland, New Zealand

[2]Adelaide Business School, University of Adelaide, Adelaide, Australia

[3]Michael Smurfit Graduate Business School, University College Dublin, Dublin, Ireland

Correspondence
Ni Yang, WF Building, 42 Wakefield Street, Auckland University of Technology, Auckland, New Zealand.
Email: ni.yang@aut.ac.nz

Abstract

We examine the mechanism by which social media sentiment affects stock prices. Specifically, we assess the impact of Twitter feeds on stock returns at the intraday level. We find that an increase in buyer-initiated trades has a significantly positive price impact. This impact, however, is stronger with an increase in the number of tweets and sentiment, and persists even after controlling for volatility, liquidity shock, and limit-order activity. The impact of Twitter sentiment on prices causes a lingering mispricing effect that is not fully assimilated at the intraday level. Rather, this mispricing takes several days to correct.

KEYWORDS
investor sentiment, return predictability, text classification, Twitter

JEL CLASSIFICATION
G11, G12, G14

1 | INTRODUCTION

Mass media outlets such as newspaper articles, opinion columns, and internet message boards affect the financial markets (Antweiler & Frank, 2004; Barber & Odean, 2008; Dougal et al., 2012; Tetlock, 2007). In recent years, however, social media has become one of the main communication channels. Social media websites such as Seeking Alpha, Motley Fool, and social media applications such as Facebook and Twitter (now rebranded as *X*) have dominated the ways in which people obtain and exchange information (Gan et al., 2020; Rakowski et al., 2021; Siganos et al., 2014). These social media applications have the advantage of disseminating information at a much higher speed than the traditional mass media outlets. Following this trend, the U.S. Securities Exchange Commission (SEC) recognizes social media as an official news announcement channel (SEC, 2013).[1]

With social media being a fast and more efficient information dissemination channel, investors can quickly update their information set and revise their trading decisions. This may affect the financial markets in several ways. First, social media posts may induce a temporary price pressure that moves prices away from the fundamentals (Baker & Wurgler, 2006; Behrendt & Schmidt, 2018; Tetlock, 2007). Second, such posts may contain information that moves prices permanently toward a new equilibrium (Azar & Lo, 2016; Bollen et al., 2011; Gu & Kurov, 2020). To disentangle the above relations, it is necessary to study the connection between social media sentiment and security prices in a high-frequency setting. In the current study, we analyze the impact of social media sentiment on market prices at the intraday level, with a particular focus on delving into its underlying mechanisms.

There is a fast-growing literature that examines the linkage between social media and security prices at the intraday level. Many of them, however, mainly explore the predictive power of social media sentiment on returns. Sun et al. (2016), for instance, study the impact of online news and media sentiment on S&P500 index returns. They find that intraday returns are predictable using lagged half-hour investor sentiment. Renault (2017) uses the social media platform *StockTwits*[2] to test the intraday sentiment effect and finds that online investor sentiment helps forecast intraday stock index returns. In particular, the first half-hour change in investor sentiment predicts the last half-hour S&P 500 index ETF return. At the firm level, Broadstock and Zhang (2019) find firms' price dynamics are susceptible to social-media sentiment pricing factors while Schnaubelt et al. (2020) show that machine learning can be used to extract predictive information from tweets that can be translated to statistically and economically significant excess returns. While these studies show that social media and markets are related, the question remains on how social media sentiment transmits to the movements in security prices. Our study fills this gap in the literature by examining the mechanism that links sentiment and prices using high-frequency data.

To study social media sentiment, we use Twitter feeds (also called "tweets") to construct a proxy for social media sentiment, which we refer to as *Twitter sentiment* hereafter. Specifically, we extract the message volume (total number of tweets), bullishness (positivity of the tweets), and agreement index (agreement about the tone of the tweets) from Twitter feeds. This choice is based on several reasons. First, Twitter is a leading social networking and microblogging service globally and has been at the forefront of online forums among investors and other finance professionals (Sprenger, Tumasjan et al., 2014). In 2020 and 2021, there were more than 500 million tweets per day (Rakowski et al., 2021). This translates to a very granular dataset, with each tweet being time-stamped to the nearest second, allowing a study at the intraday level. Second, unlike other social media platforms with postings of images and videos, Twitter specializes in instant messages with a cap of 280 characters. The textual nature of tweets allows for the extraction of social media sentiment. It overcomes many issues related to an indirect data source, such as answering bias (e.g., survey), the idiosyncratic non-sentiment-related indicator[3] (e.g., trading volume, option implied volatilities), or confounding causality (e.g., Google search queries). Third, historical tweets data are reliable and have been used in many studies (see, e.g., Azar & Lo, 2016; Rakowski et al., 2021; Schnaubelt et al., 2020).

We focus on tweets related to the SPDR S&P 500 ETF (ticker symbol: SPY) as a representation of the U.S. stock market. Baker and Wurgler (2007) define investor sentiment as a belief about future cash flows and investment risks that are not justified by economic fundamentals. As such, sentiment may not be derived from financial market conditions, and therefore, exogenous to economic fundamentals. Therefore, to assess the impact of social media sentiment on stock prices, we rely on the assumption that investor sentiment is an exogenous shock to the financial markets (see, e.g., Baker & Wurgler, 2007; Behrendt & Schmidt, 2018; Siganos et al., 2014). This assumption is common in the

[1] See, https://www.sec.gov/news/press-release/2013-2013-51htm.

[2] StockTwits is a social microblogging platform dedicated to financial markets where users share information about the market and individual stocks, in the form of short messages.

[3] This refers to the indicator being more a proxy of equilibrium of economic forces than investor sentiment (Baker & Wurgler, 2007).

literature.[4] Siganos et al. (2014) use Facebook gross national happiness index as a proxy of sentiment and find that sentiment on Sunday affects stock returns on Monday, suggesting causality from sentiment to stock markets. Based on this assumption, we modify Hasbrouck's (1991) informativeness of trades model by interacting the sentiment measure with the trade variable in the vector autoregression (VAR) model.[5] This modified VAR model allows us to measure the price impact of social media sentiment through its impact on trades. In the robustness section, we alleviate this exogeneity assumption through an instrumental variable analysis, reaching similar conclusions.

We find that Twitter sentiment intensifies the impact of trades on returns. Trades have a greater price impact with an increase in the number of tweets or a higher degree of bullishness. This finding suggests that as there is more social media discussion related to the stock market, the price impact of trades becomes stronger. Similarly, the impact of buyer- (seller-)initiated trades on stock prices is stronger when the market sentiment is bullish (bearish). This price impact persists even after controlling for volatility, liquidity shocks, and the effect of quotes on returns. Furthermore, both bullish and bearish tweets amplify the effect of trades on returns, indicating no asymmetric price impact of Twitter sentiment. Finally, this impact of Twitter sentiment on prices causes a lingering mispricing effect that is not fully assimilated at the intraday level. Rather, we find that Twitter sentiment shock on SPY returns lasts approximately 2 days, after which the mispricing is fully corrected. These findings suggest that, at the intraday level, market continues to incorporate social media sentiment. However, over the subsequent days, the impact of sentiment diminishes, leading to a reversal in prices. This underscores that the initial price response to a sentiment shock was likely attributed to noise, and it takes several days for the correction to take place.

Our study contributes to the literature in several ways. First, we add to the understanding of the impact of social media sentiment at the intraday level (see, e.g., Broadstock & Zhang, 2019; Renault, 2017; Sun et al., 2016). By modelling the channel by which sentiment affects stock price movements, our findings reveal a significant impact from social media sentiment to stock prices through its impact on trades at the intraday level, a mechanism that, to our best knowledge, has not previously been documented in the literature. Second, by examining whether social media sentiment temporarily or permanently affects prices, we contribute to the ongoing debate on whether social media sentiment provides a noisy or an informative signal (see, e.g., Karagozoglu & Fabozzi, 2017; Schnaubelt et al., 2020; Sprenger, Sandner et al., 2014). Our findings support the view that sentiment exerts a temporary price pressure that lasts for days. However, at the intraday level, the market takes time to incorporate social media sentiment, presenting a discernible intraday pattern that can be effectively exploited by day traders.

The closest papers to ours are Kurov (2008) and Kordonis et al. (2016). Kurov (2008) examines the impact of investor sentiment on the trading behaviors of index futures traders. He finds that higher investor sentiment leads to more active trading, and that order flow is less informative when investors are optimistic. Kordonis et al. (2016) analyze and extract sentiment from daily Twitter tweets. Their primary objective was to use sentiment to forecast daily stock prices of several technology stocks, that is, the top 16 technological stocks listed on Yahoo! Finance. Our study differs from them in several aspects. First, unlike Kordonis et al. (2016) who utilize- sentiment for forecasting purposes, our focus is on demonstrating the channel in which sentiment impacts stock returns at a high frequency level. To do so, we draw on insights from the market microstructure literature. More precisely, we adapt the market microstructure model proposed by Hasbrouck (1991), employing a modification akin to that employed by Dufour and Engle (2000). More specifically, we interact social media sentiment with the stock market order flow and examine how this joint variable affects stock returns. This modification allows us to examine the channel by which sentiment contributes to the intraday price formation.

Second, existing studies tend to use low frequency data to study social media sentiment. Kurov (2008), for instance, employs weekly sentiment data to study the impact of sentiment on high-frequency stock prices data (at a 1-min

[4] According to De Long et al. (1990), rational arbitrageurs are sentiment-free traders, while irrational traders are more apt to exogenous sentiment. Studies have also documented that investor sentiment could be driven by exogenous factors such as weather (Goetzmann et al., 2015; Hirshleifer & Shumway, 2003), seasonal affective disorders (Kamstra et al., 2003), day-of-the-week (Birru, 2018; Hirshleifer et al., 2020), daylight saving change (Kamstra et al., 2000), or sporting events (Edmans et al., 2007).

[5] Similar modification is applied in Dufour and Engle (2000) where they assess the price impact of trade duration on security prices.

frequency).[6] This difference in data frequency creates a non-synchronicity issue between prices and the sentiment measure, that is, the intraday market reaction is matched with investor sentiment over the week. Kordonis et al. (2016), although also using Twitter data, perform their study by aggregating tweets at a daily frequency. Our study, on the other hand, utilizes intraday frequency constructed from data at the tick level and Twitter data at the 1-min level. This difference is crucial because studies on low frequency sentiment generally show that the impact of sentiment is transitory, that is, a positive sentiment leads to positive price changes which will then correct after a few days. Gao et al. (2020), for instance, observe a reversal sentiment effect after three days on average. In contrast, we show that at the intraday horizon, financial markets can exhibit inefficiency, resulting in price changes that may not necessarily revert within the course of the day. Finally, Kordonis et al. (2016) focus on the stocks of a few big high-tech companies while our study is at the market level, that is, the SPDR S&P 500 Trust ETF (ticker: SPY) as a representation of the US stock market. As such, our findings are more generalizable.

The remainder of the paper proceeds as follows. Section 2 introduces data sources and how we construct Twitter sentiment. In Section 3, we present the methodology. We report the empirical results in Section 4 and the robustness tests in Section 5. Section 6 concludes.

2 | DATA

2.1 | Twitter sentiment

We focus on the SPDR S&P 500 Trust ETF (ticker: SPY) as a representation of the US stock market. We collect tweets from Twitter. Following Sprenger, Tumasjan et al. (2014), we use *cashtags*, denoted by the "$" sign, to specifically search for tweets pertaining to a particular security. Distinct from *hashtags* that encapsulate words and topics (denoted by the "#" sign), cashtags contain the ticker symbol of a security, that is, "$SPY" to obtain tweets related to SPY. Since Twitter officially introduced cashtags only in July 2012, we focus our sample period from August 1, 2012, to December 31, 2018, a total of 1610 trading days. There are 2.18 million tweets related to $SPY in total. Every tweet is reported in Eastern Standard Time (EST) and timestamped to the nearest second. We focus on tweets during the trading hours between 9:35 and 15:55 EST to match our stock market data.[7] Each tweet is cleaned from irrelevant characters, including emojis and links.

Similar to Azar and Lo (2016), we use an extensive electronic lexical database as a language processing tool to assign a tone score for each tweet in order to group it into negative, neutral, or positive categories (ranging from −1 to +1). More specifically, we use a lexical database *WordNet*,[8] which is widely adopted for social media sentiment evaluation and classification (see, e.g., AlMousa et al., 2021; Bhala & Abirami, 2014; Kocoń & Maziarz, 2021; Navigli, 2009). We access WordNet via a Python package *TextBlob*, which provides an Application Programming Interface (API) for natural language processing, including phrase extraction, sentiment analysis, and classification. We consider a tweet as positive if its tone score is positive, negative if the score is negative, and neutral when it is zero. In the robustness section, we show that the main results do not hinge on how we define the negative, neutral and positive tweets.

To assess the accuracy of our method, we follow Antweiler and Frank (2004) and manually conduct a pilot test on 1000 tweets and group them into negative, neutral, or positive categories before applying the dictionary to the entire

[6] Kurov (2008) uses two sentiment measures based on weekly surveys conducted by *the Investor Intelligence* (representing the outlook of about 150 independent market newsletters) and the American Association of Individual Investors (AAII), respectively. Sentiment measure based on survey data suffers several limitations including susceptibility to answering bias, information distortion, and unavoidable "cross-validation" limitations (see, e.g., Conrad et al., 2021; Schober et al., 2016; Vosen & Schmidt, 2011).

[7] We exclude the first and last 5 min of trading to minimize the confounding effects of market opening and closing.

[8] This database contains 155,327 nouns, verbs, adjectives, and adverbs organized in 175,979 synsets (sets of one or more synonyms that are interchangeable in some context) of cognitive synonyms for a total of 207,016 word-sense pairs of English for sentiment analysis. Additionally, all synsets are interlinked through conceptual-semantic and lexical relations as the computational lexicon of English for word sense disambiguation (Navigli, 2009).

dataset. Our method provides the correct groupings in approximately 90% of the training dataset. This is comparable to Antweiler and Frank (2004) who obtain an accuracy of 88% for their training dataset.

We aggregate the tweets data to a minute frequency to ensure a continuous series. We then use these aggregated data to construct three sentiment measures commonly used in the literature (see, e.g., Antweiler & Frank, 2004; Sprenger, Tumasjan et al., 2014), which we generally call *Twitter sentiment*. First, to assess whether the total number of tweets impacts stock market prices, we construct $MessageVol_t$ as follows,

$$MessageVol_t = \ln\left(1 + M_t^{Positive} + M_t^{Negative} + M_t^{Neutral}\right), \tag{1}$$

where $M_t^{Positive}$, $M_t^{Negative}$, and $M_t^{Neutral}$ are the sum of positive, negative, and neutral tweets in the minute interval t, respectively. Second, we construct a bullishness signal, $Bullishness_t$, as follows,

$$Bullishness_t = \ln\left[\frac{1 + M_t^{Positive}}{1 + M_t^{Negative}}\right]. \tag{2}$$

This measure captures the sentiment embedded in tweets during the minute interval t. Positive (negative) bullishness reflects positive (negative) sentiment. Third, we compute $AgreementIndex_t$ to measure the prevailing level of agreement among tweets as,

$$AgreementIndex_t = 1 - \sqrt{1 - \left(\frac{M_t^{Positive} - M_t^{Negative}}{M_t^{Positive} + M_t^{Negative}}\right)^2}. \tag{3}$$

This index ranges from 0 and 1, where a higher value indicates greater agreement.

Panel A of Table 1 reports daily summary statistics for tweets data. On average, there are 703 tweets related to SPY per day during the trading hours or roughly two tweets per minute. On average, tweets are slightly positive about the SPY, with a positive average bullishness of 0.16. This result is consistent with the existing literature that investors are generally more optimistic in the markets (Baker & Stein, 2004; Kim & Kim, 2014; Miller, 1977; Stambaugh et al., 2012). The average AgreementIndex is 0.42 with a slight variation in its percentiles from 0.31 (5^{th} percentile) to 0.52 (95^{th} percentile). This indicates that, on average, there is more disagreement among tweets related to the stock market.

2.2 | Stock market data

For the stock market data, we collect transaction-level data of SPY during the market trading hours between 9:35 and 15:55 EST over the sample period from Refinitiv Tick History. The data contains all activity observed at the national best bid and offer, which includes recorded transactions and revisions in the bid and ask prices and depths, all time-stamped to the nearest millisecond. Our observations from the intraday data show a number of anomalous records that appear to be recording errors. Therefore, we remove transactions where trading volume is above the day's 99.9^{th} percentile. We then follow Chordia et al. (2001) and drop observations using the following filters: (1) non-positive quoted spread; (2) quoted spread greater than 5; (3) effective spread/quoted spread greater than 4; (4) percentage effective spread/percentage quoted spread greater than 4; (5) quoted spread/transaction price greater than 0.4.

We treat multiple trades that are executed with the same timestamp as one trade, as they typically reflect a trade initiated by one market participant but executed against the limit orders of multiple market participants. In such cases, we use the value-weighted average price and aggregate the volume traded. Each trade is classified into buyer- and seller-initiated trades using the Lee and Ready (1991) algorithm. A trade is classified as buyer- (seller-) initiated if the transaction price is above (below) the prevailing midquote. For trades that occur at the midquote, we employ the tick

TABLE 1 Summary statistics.

	Mean	S.D.	Median	5th Percentile	95th Percentile	AR(1)
Panel A: Tweets						
TotalTweets	703	348	621	350.45	1,293.55	0.74
MessageVol	0.84	0.25	0.79	0.51	1.29	0.71
Bullishness	0.16	0.07	0.15	0.07	0.31	0.51
AgreementIndex	0.42	0.07	0.43	0.31	0.52	0.56
Panel B: SPY						
Trades('000)	126	70	106	64	256	0.79
Volume('000)	71,170	32,848	65,169	33,120	1,320,718	0.71
OIBV('000)	196	2163	137	−3265	3806	0.10
MidReturn(bps)	1.07	65.60	4.38	−109.56	92.28	−0.04

Panel C: Pairwise correlations

	TotalTweets	*MessageVol*	*Bullishness*	*AgreementIndex*	*Trades*	*Volume*	*OIBV*	*MidReturn*
TotalTweets	1.00							
	–							
MessageVol	0.96***	1.00						
	(140.28)	–						
Bullishness	0.60***	0.63***	1.00					
	(30.31)	(32.74)	–					
AgreementIndex	0.60***	0.75***	0.65***	1.00				
	(30.19)	(45.19)	(34.71)	–				
Trades	0.73***	0.76***	0.45***	0.50***	1.00			
	(42.59)	(46.99)	(20.20)	(23.16)	–			
Volume	0.52***	0.54***	0.23***	0.34***	0.76***	1.00		
	(24.59)	(26.02)	(9.42)	(14.63)	(47.30)	–		
OIBV	−0.09***	−0.10***	0.02	−0.06**	−0.03	0.05	1.00	
	(−3.74)	(−3.91)	(0.73)	(−2.29)	(−1.37)	(1.41)	–	
MidReturn	−0.18***	−0.16***	0.08***	−0.02	−0.15***	−0.15***	0.40***	1.00
	(−7.31)	(−6.40)	(3.08)	(−0.93)	(−6.13)	(−6.17)	(17.49)	–

Note: This table reports the daily summary statistics for the tweets (Panel A), the SPY market data (Panel B), and the pairwise correlation coefficients (Panel C) for the sample period August 1, 2012 to December 31, 2018 (1610 trading days) during the trading hours from 9:35 to 15:55 ET. *TotalTweets* is the total number of positive, negative, and neutral tweets. *MessageVol* is the logarithm of the total number of tweets. *Bullishness* signals investors' sentiment. *AgreementIndex* indicates the degree of agreement among tweets. *Trades* is the total number of trades. *Volume* is the total trading volume. *OIBV* is the signed volume order flow. *MidReturn* is the daily return computed using intraday midquotes. S.D. is the standard deviation. Newey-West corrected *t*-statistics are in parentheses.

rule and compare the current with the previous transaction price, that is, buyer- (seller-) initiated if transaction price is above (below) the previous transaction price and undetermined otherwise. Similar to the tweets data, we aggregate the transaction-level data at a minute interval. We then compute the aggregate signed trading volume for each interval. Finally, to reduce the effect of outliers, we winsorize the minute level data per day at 2.5% each tail.

Panel B of Table 1 reports the daily statistics for SPY over the sample period. We report the daily number of trades, trading volume, signed order flow (OIBV), and midquote return. On average, there are 126,000 trades and 71 million ETF units traded each day. In terms of order flow, there are more buy transactions on average, as shown by the positive figure. The average daily midquote return is 1.07 basis points (bps).

Panel C reports the pairwise correlation matrix between tweets and market data with data aggregated at a daily frequency. In general, our sentiment measures (*MessageVol*, *Bullishness* and *AgreementIndex*) are highly correlated with each other with correlation coefficients ranging from 0.63 (between *MessageVol* and *Bullishness*) to 0.75 (between *MessageVol* and *AgreementIndex*). Sentiment is positively correlated with *Trades* and *Volume*, indicating that periods of high sentiment are associated with days of high trading activity in the equity markets, in line with the observation of Baker and Stein (2004). Both *MessageVol* and *AgreementIndex* are associated with lower signed order flow, *OIBV*. At the daily level, the linkages between our sentiment measures and *MidReturn* are mixed, that is, positive and significant for *Bullishness*, negative for *MessageVol* and insignificant for *AgreementIndex*. In the next section, we explore these linkages further at the intraday level.

3 | METHODOLOGY

To investigate the price impact of social media sentiment on stock market, we build on Hasbrouck's (1991) informativeness of trade model. Hasbrouck jointly models the generating processes of quotes and trades using an l-lags vector autoregressive (VAR) specification as follows,

$$r_k = \sum_{i=1}^{l} \alpha_i r_{k-i} + \sum_{i=0}^{l} \beta_i x_{k-i} + \varepsilon_{1,k},$$

$$x_k = \sum_{i=1}^{l} \mu_i r_{k-i} + \sum_{i=1}^{l} \delta_i x_{k-i} + \varepsilon_{2,k}, \quad (4)$$

where r_k represents the midquote return at trade k and x_k is a trade indicator that equals 1 for buyer-initiated and −1 for seller-initiated trades. The terms $\varepsilon_{1,k}$ and $\varepsilon_{2,k}$ are mutually and serially uncorrelated white noises that represent trade-unrelated and trade-related shocks, respectively. Hasbrouck (1991) imposes that in the first equation, both the contemporaneous and lagged trades can impact returns, but in the second equation, only the lagged returns can impact trades. This model is estimated using ordinary least squares (OLS).

To understand the role of sentiment on movements in prices, we modify the VAR model in Equation (4) by including a Twitter sentiment measure. More specifically, we interact the trade variable, x, in the VAR with a measure of Twitter sentiment. Following the literature (Baker & Wurgler, 2007; Behrendt & Schmidt, 2018; Siganos et al., 2014), we assume that sentiment is an exogenous variable in the VAR system, that is, sentiment could affect trades and returns, but not the other way round.[9] Unlike Hasbrouck (1991) who examines the price impact after each trade, we replace the trade variable, x, in Equation (4) with the signed volume order flow for each minute interval t.[10] Time aggregation of returns and order flow reduce the synchronicity issue between the trade prices and trade sizes (Kurov, 2008). At the same time, it allows us to ensure a continuous time series for our tweets and market data. Furthermore, Chordia et al. (2005) explain that signed volume order flow is a better proxy for trading activity and has a more meaningful relation to the direction and magnitude of price changes. Thus, our modified VAR model is as follows,

$$r_t = \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \theta_i STMT_{t-i}) x_{t-i} + \varepsilon_{1,t},$$

[9] To alleviate this exogeneity assumption, in the robustness Section 5.1 we implement an instrumental analysis using the lagged Twitter sentiment variables as an instrument.

[10] Brown et al. (1997), Kurov (2008), and Huang et al. (2021) also use time aggregation for the signed order flow to replace the trade variable in Hasbrouck's (1991) VAR model.

$$x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \lambda_i STMT_{t-i}) x_{t-i} + \varepsilon_{2,t}, \quad (5)$$

where $STMT_t$ is either the *MessageVol*, *Bullishness*, or *AgreementIndex* for each minute interval t, and x_t is the signed volume order flow. Following Dufour and Engle (2000), we employ five lags in the VAR. The coefficients of interest are θ_i's, which are the coefficients of the interactions $STMT_{t-i} \times x_{t-i}$ in the return equation. These coefficients provide evidence of the impact of Twitter sentiment on prices at high frequency. If $\sum_{i=0}^{5} \theta_i$ is significantly different from zero, we can conclude that the Twitter sentiment affects SPY prices through its impact on trades. In addition, we compare the sign of the coefficients β_i and θ_i in the return equation. If both coefficients have the same (different) signs, it suggests that Twitter sentiment enhances (weakens) the impact of trades on prices.

4 | EMPIRICAL RESULTS

4.1 | Price impact of tweets

To investigate the price impact of Twitter sentiment, we estimate Equation (5) for each day using OLS based on data aggregated at the 1-min frequency. To ease comparison, we normalize all variables in Equation (5). In Table 2, we report the average of those estimates along with their Newey-West corrected t-statistics.

Turning first to Panel A, we find that the three different sentiment measures provide similar results. For instance, the sum of α_i coefficients for *MessageVol* is −0.0736 and statistically significant at the 1% level. This indicates a mean reversion in returns. The sum of β_i coefficients is 0.4333, also statistically significant at the 1% level. This implies that midquote returns increase (decrease) with buy (sell) orders, consistent with Hasbrouck's (1991).

We are particularly interested in the coefficients θ_i's, that is, the coefficients of the interaction term $STMT_{t-i} \times x_{t-i}$ in the return equation. The coefficient for the contemporaneous effect θ_0 is positive and statistically significant for all three sentiment measures, indicating that the impact of sentiment on prices occurs relatively quickly. The sum of θ_i coefficients is positive and statistically significant at the 1% level for message volume and at the 5% level for bullishness. Our findings indicate that the impact of trades on SPY prices is affected by the increase in the number of tweets and the bullish sentiment on Twitter about SPY. Given that β_i and θ_i share the same sign, our results also indicate that the impact of trades on intraday returns is stronger when it is interacted with Twitter sentiment. In other words, Twitter sentiment enhances the effect of trades on prices. This finding extends the results of Antweiler and Frank (2004), Sprenger, Tumasjan et al. (2014), and Schnaubelt et al. (2020) by showing that sentiment influences intraday stock market returns through its impact on trades. Similarly, we find that *AgreementIndex* has a positive marginal effect on the price impact of trades. However, we do not find this effect to be statistically significant, indicating that returns are not significantly affected by the divergence level of investors' opinions on Twitter. This finding suggests that the agreement index is conceptually different from bullishness and message volume.

Panel B of Table 2 shows the impact of Twitter sentiment on the signed volume order flow. The sum of δ_i coefficients for the signed volume order flow is positive and statistically significant at the 1% level. This indicates persistence in order flow, that is, purchases tend to follow purchases, and sales tend to follow sales. Examining the λ_i coefficients for the interaction terms, we observe that they are negative for all three sentiment measurements, although they are only statistically significant for the message volume. We interpret this finding as a higher message volume reduces the persistence in trades in the subsequent minutes. In other words, when there are more tweets about SPY, the persistence in order flow will be less pronounced. We do not observe this persistence to be affected by either the level of Bullishness or AgreementIndex.

In summary, Twitter sentiment is linked to intraday stock returns. In particular, the increase in volume and bullishness of tweets can intensify the impact of trades on intraday returns. Unlike the first two, however, AgreementIndex only affects returns contemporaneously. These findings indicate that social media sentiment is transmitted to stock

TABLE 2 Vector autoregression estimation for Twitter sentiment.

	α_i		β_i		θ_i								
Panel A: Return equation	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Lag0*	*Lag1*	*Lag2*	*Lag3*	*Lag4*	*Lag5*	*Sum of lags*	*t-stat*	*Adj-R*^{2}
MessageVol	−0.0736***	(−17.49)	0.4333***	(66.06)	0.0185***	0.0043***	0.0022*	0.0016	0.0016	0.0014	0.0297***	(10.43)	0.33
Bullishness	−0.0758***	(−17.49)	0.4385***	(64.64)	0.0027*	0.0022*	−0.0004	0.0003	−0.0007	0.0025**	0.0064**	(2.08)	0.33
AgreementIndex	−0.0748***	(−17.29)	0.4382***	(67.78)	0.0051***	−0.0010	−0.0022*	0.0003	0.0008	0.0006	0.0036	(1.16)	0.33

	μ_i		δ_i		λ_i							
Panel B: Trade equation	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Lag1*	*Lag2*	*Lag3*	*Lag4*	*Lag5*	*Sum of lags*	*t-stat*	*Adj-R*^{2}
MessageVol	−0.0043	(−0.46)	0.2512***	(68.49)	−0.0086***	−0.0086***	−0.0030**	−0.0038**	−0.0009	−0.0249***	(−8.50)	0.05
Bullishness	−0.0050	(−0.54)	0.2493***	(70.63)	−0.0016	0.0000	−0.0010	−0.0006	0.0017	−0.0016	(−0.50)	0.05
AgreementIndex	−0.0057	(−0.62)	0.2492***	(69.93)	−0.0018	−0.0001	−0.0008	0.0000	0.0003	−0.0024	(−0.80)	0.05

Note: This table reports the coefficient estimates of the VAR model for Twitter sentiment impacts. The estimation model is as below, the results of return equation are presented in Panel A, and the results of trade equation are presented in Panel B.

$$r_t = \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \theta_i STMT_{t-i}) x_{t-i} + \varepsilon_{1,t}$$
$$x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \lambda_i STMT_{t-i}) x_{t-i} + \varepsilon_{2,t}$$

where x_t is volume of signed order flow at the time interval t, r_t is trade return calculated from the natural logarithm of midquote at the t-th min, and $STMT_{t-i}$ is one of the following Twitter sentiment measures: (i) *MessageVol*: the logarithm of the total number of tweets, (ii) *Bullishness*: a measure of investors' sentiment, or (iii) *AgreementIndex*: a measure of agreement among tweets. All variables are normalized. Newey-West corrected *t*-statistics are in parentheses. ***, **, and * indicate 1%, 5%, and 10% significance level.

prices through its impact on trades at the intraday level. In the next section, we further examine the informativeness of Twitter sentiment.

4.2 | Does Twitter sentiment convey information or noise?

So far, we have shown that Twitter sentiment has a significant impact on stock returns at the intraday level. Specifically, the impact of trades on returns is stronger when Twitter sentiment (message volume and bullishness) is high. However, an important question is whether Twitter sentiment captures information or noise. To answer this question, we estimate the impulse response of the SPY returns to one shock on Twitter sentiment using the local projections approach of Jordà (2005). The analysis of local projections is an intuitive and simple approach (see the local projections survey of Jordà, 2023), as it involves conducting OLS regressions across multiple horizons, h. Additionally, by applying Newey-West corrected standard errors, it accurately estimates the impulse responses to a single shock in the sentiment variable. That is, we run the following regression by OLS,

$$r_{t+h} = \sum_{i=1}^{5} \alpha_i r_{t-i} + \beta_0 x_t + \theta_{0,h} STMT_t x_t + \sum_{i=1}^{5} (\beta_i + \theta_i STMT_{t-i}) x_{t-i} + \varepsilon_{h,t}. \quad (6)$$

Note, this equation is similar to the return equation in Equation (5) with the only difference being the independent variable is now r_{t+h}, which is the midquote return at h minute ahead with $h = \{0, 1, 2, \ldots\}$. The coefficient of interest is $\theta_{0,h}$ which captures the impact of the contemporaneous interaction term between the Twitter sentiment and OIBV. The impulse response at horizon h is calculated as $IRF(t, h, STMT) = \hat{\theta}_{0,h} \sigma_{STMT} \bar{x}$, where $\hat{\theta}_{0,h}$ is the estimated contemporaneous coefficient of Equation (6), σ_{STMT} is the standard deviation of the Twitter sentiment measure, and $\bar{x}$ is the average of OIBV at day t. Similar to Equation (5), we normalize all the variables in Equation (6), gauge the cumulative impulse response for each day of our sample and plot its average and 95% confidence bands in Figure 1.

The impulse response based on local projections in Figure 1 reveals a noteworthy trend: following an intraday positive message volume or bullishness shock, prices increase and stabilize within the first hour. This pattern may suggest that Twitter sentiment carries informative value regarding the fundamentals of SPY, indicating a potentially permanent influence of social media on prices.[11]

While a lasting price movement due to Bullishness and Message Volume may indicate the integration of new information and the establishment of a new equilibrium, it is crucial to acknowledge that the lasting price movement could also be due to a lingering mispricing effect of sentiment on prices that may not have been fully assimilated at the intraday level. This observation aligns with findings from other studies such as Gao et al. (2020) and Edmans et al. (2022) who show that sentiment-induced temporary price changes often undergo correction over several days.

To test this conjecture, we examine the cumulated price impact of tweets at a daily frequency. Similar to the local projections at intraday frequency, we aggregate the three sentiment measures, the volume of signed order flow and return of SPY to daily frequency, then we re-run the impulse response by local projections in Equation (6) with daily data. We compute the cumulated return dynamics at daily frequency over our sample period and plot it in Figure 2. The lower and upper bands are the 95% confidence interval with Newey-West corrected standard errors from Equation (6).

Figure 2 reveals that, in general, the cumulative return still remains statistically significant approximately 1–2 days post a sentiment shock and becomes insignificant after that. Combined, our findings suggest that, at the intraday level, market continues to incorporate social media sentiment, especially for Bullishness. However, over the subsequent days, the impact of sentiment diminishes, leading to a reversal in prices. This evidence confirms that the initial reaction of prices to a Twitter sentiment shock was likely attributed to noise, and it takes several days for the correction to take place.

[11] We acknowledge that Agreement Index leads to an erratic price dynamic, suggesting that this variable is distinct from Bullishness and Message Volume. Nevertheless, these results are in line with Table 2 where Agreement Index where its cumulate effect was insignificant.

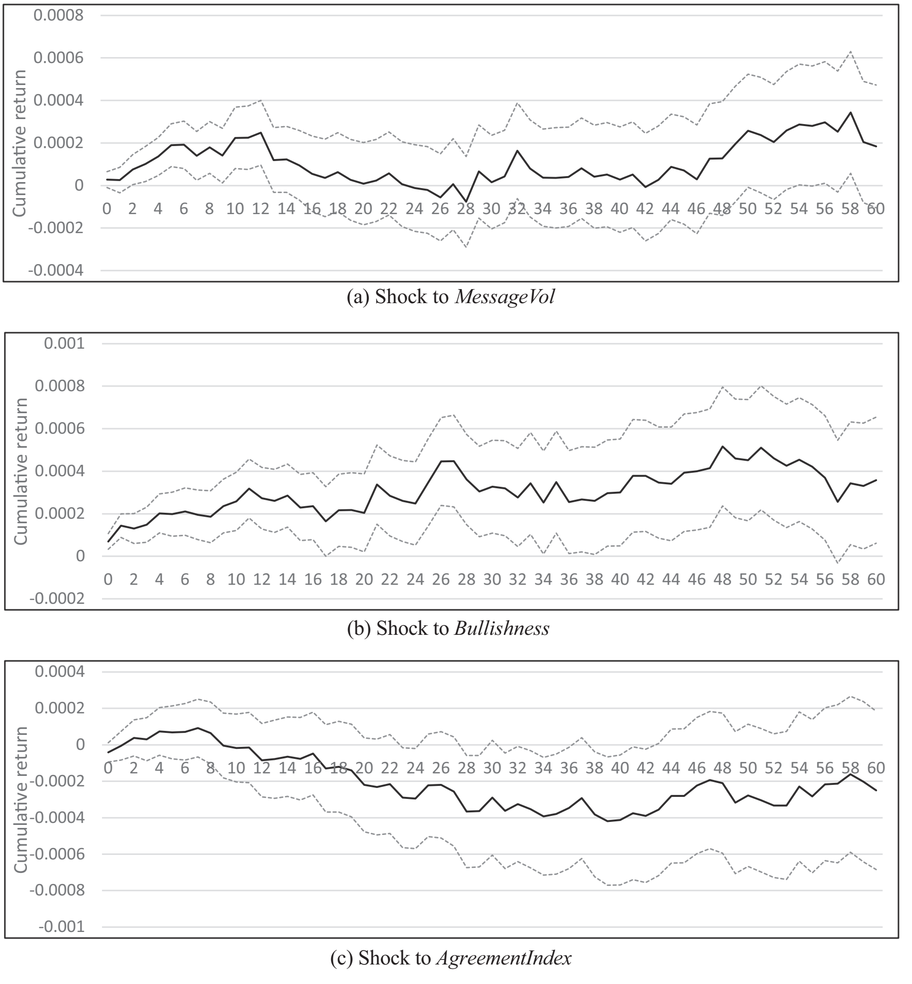


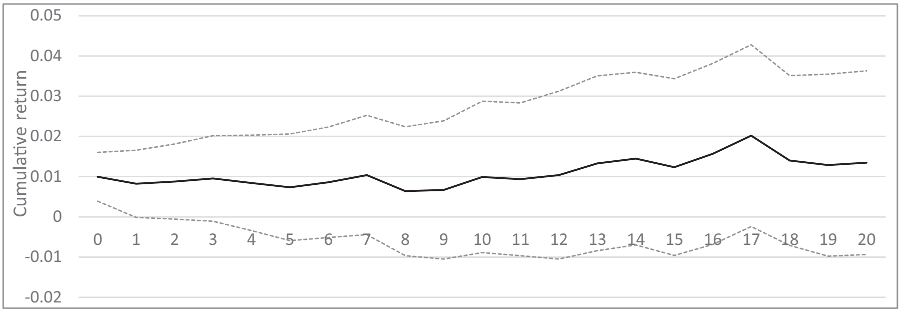
FIGURE 1 Analysis of impulse responses at an intraday frequency.
This figure plots the intraday cumulative returns following the estimation model below. It is based on the impulse response by local projections with one standard deviation shock to the Twitter sentiment measures.

$$r_{t+h} = \sum_{i=1}^{5} \alpha_i r_{t-i} + \beta_0 x_t + \theta_{0,h} STMT_t x_t + \sum_{i=1}^{5} (\beta_i + \theta_i STMT_{t-i}) x_{t-i} + \varepsilon_{h,t},$$

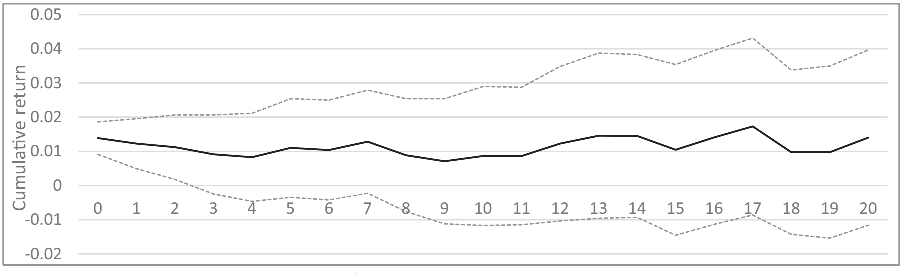
where r_{t+h} is the midquote return at h minute ahead with lags of r_t included, x_t is volume of signed order flow at the time interval t, and $STMT_{t-i}$ is one of the Twitter sentiment measures. The impulse response is calculated as $IRF(t, h, STMT) = \hat{\theta}_{0,h}\, \sigma_{STMT} \bar{x}$ where $\hat{\theta}_{0,h}$ is the estimated contemporaneous coefficient of the above equations, σ_{STMT} is the standard deviation of the sentiment measure and $\bar{x}$ is the average of OIBV at intraday frequency. All variables are normalized. The 60-minute cumulative return is the average of intraday forecast across 1610 days. The lower and upper bands are the confidence intervals at 5% significance level based on Newey-West corrected standard errors.

4.3 | The importance of positive and negative tweets

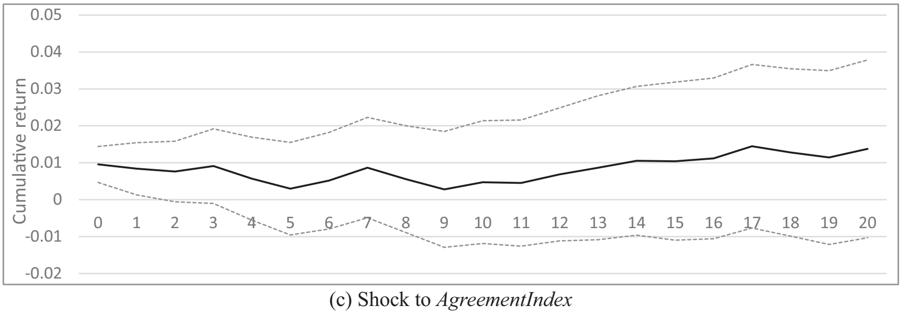
Next, we examine if the impact of Twitter sentiment on stock returns is asymmetric. Currently, the literature shows mixed evidence on whether positive or negative sentiment has a stronger impact on prices. For example, Chen et al. (2004) and Barber and Odean (2008) show that positive sentiment has a stronger impact on returns than negative



(a) Shock to *MessageVol*



(b) Shock to *Bullishness*



(c) Shock to *AgreementIndex*

FIGURE 2 Analysis of impulse responses at a daily frequency.
This figure plots the daily cumulative returns following the estimation model below. It is based on the impulse response by local projections with one standard deviation shock to the Twitter sentiment measures.

$$r_{t+h} = \sum_{i=1}^{5} \alpha_i r_{t-i} + \beta_0 x_t + \theta_{0,h} STMT_t x_t + \sum_{i=1}^{5} (\beta_i + \theta_i STMT_{t-i}) x_{t-i} + \varepsilon_{h,t},$$

where r_{t+h} is the midquote return at time h ahead with lags of r_t included, x_t is volume of signed order flow at the time interval t, and $STMT_{t-i}$ is one of the Twitter sentiment measures. The impulse response is calculated as $IRF\ (t, h, STMT) = \hat{\theta}_{0,h}\ \sigma_{STMT} \bar{x}$ where $\hat{\theta}_{0,h}$ is the estimated contemporaneous coefficient of the above equations, σ_{STMT} is the standard deviation of the sentiment measure and $\bar{x}$ is the full sample average of OIBV at daily frequency. All variables are normalized. The cumulative return is at a daily frequency. The lower and upper bands are the confidence intervals at 5% significance level based on Newey-West corrected standard errors.

sentiment, whereas Akhtar et al. (2011, 2012) and Agrawal et al. (2018) find that negative sentiment on media exerts a stronger effect on stock markets than positive sentiment.[12] Others find that the impact of positive and negative sentiments is symmetric (see, e.g., Al-Nasseri et al., 2021; Moseki & Rao, 2018).

[12] The "positive sentiment effect" can be attributed to retail investors' limited cognitive processing ability (Chen et al., 2004) and attention-grabbing nature of the stock, for example, stocks in the news, stocks experiencing high abnormal trading volume, and stocks with extreme one-day returns (Barber & Odean,

TABLE 3 Vector autoregression estimation for positive and negative Twitter sentiment.

	τ_i		ω_i		Paired coefficients test	
	Sum of lags	*t-stat*	*Sum of lags*	*t-stat*	*t-stat*	*Adj-R²*
Panel A: Return equation	0.0132***	(4.22)	0.0093***	(2.76)	(0.81)	0.34
	ζ_i		ς_i		**Paired coefficients test**	
	Sum of lags	*t-stat*	*Sum of lags*	*t-stat*	*t-stat*	*Adj-R²*
Panel B: Trade equation	−0.0078**	(−2.53)	−0.0079**	(−2.37)	(0.03)	0.05

Note: This table reports the coefficient estimates of the VAR model for positive and negative tweets. The estimation model is as below, the results for the return equation are presented in Panel A, and the results for the trade equation are presented in Panel B.

$$r_t = \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \tau_i POS_{t-i} + \omega_i NEG_{t-i}) x_{t-i} + \varepsilon_{1,t}$$
$$x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \zeta_i POS_{t-i} + \varsigma_i NEG_{t-i}) x_{t-i} + \varepsilon_{2,t}$$

where x_t is volume of signed order flow at the time interval t, r_t is trade return calculated from the natural logarithm of midquote at the t-th min. *POSt* and NEG_t are the total number of positive and negative tweets over the minute interval t, respectively. The determination of a positive (negative) tweet is based on the tone score greater (less) than 0 explained in Section 3.1. All variables are normalized. Newey-West corrected t-statistics are in parentheses. ***, **, and * indicate 1%, 5%, and 10% significance level.

Given the mixed evidence, we examine if there is an asymmetric effect from Twitter sentiment. We first classify tweets into positive and negative sentiment where POS_t and NEG_t represent the sum of all tweets with positive and negative tone scores during the minute interval t, respectively. We then modify the VAR model as follows,

$$r_t = \sum\nolimits_{i=1}^{5} \alpha_i r_{t-i} + \sum\nolimits_{i=0}^{5} (\beta_i + \tau_i POS_{t-i} + \omega_i NEG_{t-i}) x_{t-i} + \varepsilon_{1,t},$$

$$x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \zeta_i POS_{t-i} + \varsigma_i NEG_{t-i}) x_{t-i} + \varepsilon_{2,t} \quad (7)$$

Table 3 reports the sum of the regression coefficients from Equation (7). The sum of coefficients for the interaction term with positive (τ_i) and negative (ω_i) sentiments in the return equation are positive (0.0132 and 0.0093, respectively) and significant at the 1% level. This finding indicates that both sentiments affect prices. These coefficients are positive as they reflect the marginal impact of trades on prices, that is, positive (negative) order imbalance will lead to even higher (lower) stock returns with the arrival of either positive or negative social media sentiments.

We also test for asymmetry using a paired t-test for the difference in coefficients for the positive and negative sentiments. The reported t-statistic (0.81) shows that the null hypothesis of equal coefficients is not rejected. These results suggest that both positive and negative sentiments symmetrically intensify the impact of trades. This is in line with Moseki and Rao (2018) and Al-Nasseri et al. (2021) who document symmetric effects between positive and negative sentiments.

4.4 | The role of volatility and liquidity

It is possible that Twitter sentiment may be proxying for other factors such as market uncertainty and liquidity. Shu and Chang (2015), for instance, show that market uncertainty is often associated with investor sentiment, while Dumas

2008). The "negative sentiment effect" can be explained by the availability heuristics of investors, that is, the tendency to form a judgment based on what is readily brought to mind (Akhtar et al., 2011, 2012).

et al. (2009) find that optimistic anticipation of the future leads to overconfidence among investors and results in overreaction and volatility in the stock markets. Other studies also document that liquidity is often associated with investor sentiment. Liu (2015) uses liquidity measures developed by Amihud (2002) and shows that stock market is more liquid when market sentiment is higher, that is, investors are more bullish. Baker and Stein (2004) show that overconfident investors tend to underreact to the information related to either order flow or equity issuance. Thus, high liquidity signals high sentiment caused by irrational investors.

To ensure that the price reaction to Twitter sentiment is not due to market uncertainty and liquidity, we control for volatility and liquidity shocks using the VAR model below,

$$r_t = \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \theta_i STMT_{t-i} + \gamma_i VIX_{t-i} + \eta_i LiqShock_{t-i}) x_{t-i} + \varepsilon_{1,t},$$

$$x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \lambda_i STMT_{t-i} + \rho_i VIX_{t-i} + \varphi_i LiqShock_{t-i}) x_{t-i} + \varepsilon_{2,t}, \quad (8)$$

where VIX_t is the Chicago Board Options Exchange (CBOE) volatility index at minute t, downloaded from Refinitiv Tick History. We use the VIX as a proxy for the U.S. market volatility. $LiqShock_t$ is the measure for liquidity shock. To construct this measure, we start by computing Amihud (2002) illiquidity ratio,

$$Amihud_t = \frac{|r_t|}{\$Volume_t}, \quad (9)$$

where r_t is the log return of SPY at minute t, and $\$Volume_t$ is the dollar trading volume during the same interval. We then follow Bali et al. (2014) and normalize this measure. More specifically, we take the negative difference between the Amihud ratio at minute t and the average value of its past 10 min and then divide by the standard deviation of the past 10-min Amihud ratios. The normalized liquidity shocks $LiqShock_t$ is computed as follows,

$$LiqShock_t = \frac{-\left[Amihud_t - \text{Mean}\left(Amihud_{t-11,t-1}\right)\right]}{SD\left(Amihud_{t-11,t-1}\right)} \quad (10)$$

where Mean($Amihud_{t-11,t-1}$) is the average Amihud illiquidity measure for SPY over the prior 10 min and $SD(Amihud_{t-11,t-1})$ is its standard deviation.[13] By normalizing the liquidity measure, we remove the expected component, leaving the unexpected or shock component in liquidity. A positive (negative) *LiqShock* indicates an unexpected increase (decrease) in SPY liquidity.

Table 4 reports the results of the VAR in Equation (8). For brevity, we only report the sum of coefficients for interaction terms with the exogenous variables, that is, the Twitter sentiment, volatility, and liquidity. For the Twitter sentiment measures, we find that the sum of parameters θ_i for message volume and bullishness remains positive at 1% and 10% significant levels, respectively. This finding suggests that the importance of Twitter sentiment is not subsumed by market uncertainty or liquidity. For the uncertainty measure, we observe that the VIX also significantly affects SPY prices through its impact on trades, that is, during periods of high market volatility or uncertainty, the price impact of trades is stronger. This is in line with the overconfident investor argument of Dumas et al. (2009). Furthermore, we observe that the impact of liquidity shock is negative and statistically significant, as shown by the negative sum of parameters η_i. This finding suggests that periods of unexpected high liquidity weaken the price impact of trade, consistent with Baker and Stein (2004).

[13] The results are qualitatively similar when we use the previous 5 or 30 min. These results are available from the authors.

TABLE 4 Vector autoregression estimation for Twitter sentiment controlling by volatility and liquidity.

	θ_i		γ_i		η_i		
Panel A: Return equation	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Adj-R²*
MessageVol	0.0207***	(7.39)	0.0221***	(7.74)	−0.4150***	(−91.70)	0.47
Bullishness	0.0052*	(1.86)	0.0244***	(8.88)	−0.4134***	(−94.74)	0.47
AgreementIndex	0.0012	(0.45)	0.0234***	(8.25)	−0.4141***	(−91.26)	0.47
	λ_i		ρ_i		φ_i		
Panel B: Trade equation	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Adj-R²*
MessageVol	−0.0260***	(−7.96)	−0.0151***	(−4.66)	−0.0384***	(−8.32)	0.06
Bullishness	−0.0026	(−0.77)	−0.0160***	(−4.90)	−0.0400***	(−8.76)	0.06
AgreementIndex	−0.0035	(−1.03)	−0.0155***	(−4.80)	−0.0398***	(−8.84)	0.06

Note: This table reports the coefficient estimates of the VAR model for Twitter sentiment impacts controlling by stock market volatility and liquidity. The estimation model is as below, the results for the return equation are presented in Panel A, and the results the trade equation are presented in Panel B.

$$r_t = \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \theta_i STMT_{t-i} + \gamma_i VIX_{t-i} + \eta_i LiqShock_{t-i}) x_{t-i} + \varepsilon_{1,t}$$

$$x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \lambda_i STMT_{t-i} + \rho_i VIX_{t-i} + \varphi_i LiqShock_{t-i}) x_{t-i} + \varepsilon_{2,t}$$

where x_t is volume of signed order flow at the time interval t, r_t is trade return calculated from the natural logarithm of midquote at the t-th min, and $STMT_{t-i}$ is one of the Twitter sentiment measures: (i) *MessageVol*: the logarithm of the total number of tweets, (ii) *Bullishness*: a measure of investors' sentiment, or (iii) *AgreementIndex*: a measure of agreement among tweets. VIX_{t-i} is the volatility index from Chicago Board Options Exchange to represent the market volatility at the time interval t. Standardized liquidity shocks is calculated as

$$LiqShock_t = \frac{-[Amihud_t - \text{Mean}(Amihud_{t-11,t-1})]}{SD(Amihud_{t-11,t-1})},$$

where $Amihud_{t-11,t-1}$ is the average value of *Amihud* over the prior 10 min. All variables are normalized. Newey-West corrected t-statistics are in parentheses. ***, **, and * indicates 1%, 5%, and 10% significance level.

4.5 | The price impact of Twitter sentiment via limit orders

Brogaard et al. (2019) show that limit orders can play a more important role in price discovery than market orders in modern financial markets. Given that high-frequency traders are more active in submitting limit orders than market orders, limit order submissions can lead to a positive price impact. We, therefore, study whether Twitter sentiment has a price impact via limit orders as an alternative channel.[14] To do so, we further modify our VAR model and endogenize limit order activity (proxied using quote-intensity-to-trade (QIT) ratio) in the VAR system as follows,

$$\begin{aligned}
r_t &= \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \theta_i STMT_{t-i}) x_{t-i} + \sum_{i=0}^{5} (\zeta_i + \eta_i STMT_{t-i}) QIT_{t-i} + \varepsilon_{1,t}, \\
x_t &= \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \lambda_i STMT_{t-i}) x_{t-i} + \sum_{i=1}^{5} (\iota_i + \nu_i STMT_{t-i}) QIT_{t-i} + \varepsilon_{2,t}, \\
QIT_t &= \sum_{i=1}^{5} \Omega_i r_{t-i} + \sum_{i=1}^{5} (\xi_i + \pi_i STMT_{t-i}) x_{t-i} + \sum_{i=1}^{5} (\kappa_i + \psi_i STMT_{t-i}) QIT_{t-i} + \varepsilon_{3,t},
\end{aligned} \tag{11}$$

where QIT_t is the quote-intensity-to-trade ratio at minute t, defined as the total number of revisions in either bid or ask prices and depths at the best quotes of the limit order book divided by the total number of trades during that minute interval. Our assumption is that higher limit order activity leads to greater revisions in prices.

[14] According to Goettler et al. (2009), Hoffmann (2014), and Roşu (2019), high-frequency traders tend to place market orders for immediate execution when they possess valuable information and place limit orders at other times.

Table 5 reports the results of the VAR in Equation (11). We focus on Panel A, which shows the impact of sentiment and limit order activity on returns. The sum of the coefficients θ_i for the interaction term $STMT \times x$ is positive and highly significant at the 1% level across all three Twitter sentiment measures. This finding suggests that, after controlling for the activity in the limit order book, the impact of Twitter sentiment on returns becomes stronger. Consistent with Brogaard et al. (2019), the sum of the coefficients ζ_i for QIT shows that more activity in the limit order book positively affects the SPY returns, showing that limit orders contain information that facilitates price discovery. However, none of the aggregated coefficients η_i for the interaction term $STMT \times QIT$ is statistically significant, indicating that Twitter sentiment does not influence returns via limit order activity.

In Panel B, we do not observe any consistent significant effect of QIT on trades, except for the negative relationship for message volume. For the QIT equation in Panel C, we observe that both higher midquote returns and positive signed order flow reduce the quote-intensity-to-trade ratio. However, the sum of coefficients π_i and ψ_i for the interaction terms of trades and QIT, respectively, are positive for each sentiment measure and significant at 1% significance level. This observation indicates that Twitter sentiment can mitigate the negative impact of trades on QIT, but strengthen the QIT autocorrelation effects. Overall, Table 5 confirms that the impact of Twitter sentiment on SPY prices is through trades, not limit order activity.

5 | ROBUSTNESS TESTS

To ensure the robustness of our results, we perform several tests. First, we address the potential endogeneity concerns of sentiment using an instrumental analysis. Second, we use a different threshold to capture extreme positive and negative tweets. Third, we consider VAR models with a different set of lags. Finally, we employ a different lexical database to classify tweets into positive, negative, and neutral.

5.1 | Further test on the exogeneity of sentiment

Our results so far show that Twitter sentiment affects stock market returns through its impact on trades. However, this is based on the assumption of sentiment being exogenous. More specifically, the VAR in Equation (5) was inspired by Dufour and Engle (2000)'s model where the exogeneity of the sentiment is one of the key assumptions. This assumption is in line with extensive literature which assumes that investor sentiment is an exogenous shock to the financial markets (see, e.g., Baker & Wurgler, 2007; Behrendt & Schmidt, 2018; Siganos et al., 2014). For example, Behrendt and Schmidt (2018) use Twitter sentiment as an exogenous variable to forecast intraday volatility of individual stocks. They find that Twitter sentiment does not significantly enhance the accuracy of the forecasts.

There is potentially a counterargument that sentiment is endogenous and caused by stock returns. For example, if the market is going up (down), investors may become more optimistic (pessimistic) and start buying (selling) stocks. To alleviate this concern, we implement an instrumental analysis. Given the difficulty of finding an instrument at intraday frequency that is correlated with Twitter sentiment, but uncorrelated with the error term in the VAR equations, we make use of the lagged intraday sentiment measures as a valid instrument. This assumption is based on the high persistence of the sentiment measures (the AR(1) coefficient in Table 1). Therefore, we modify the VAR system in Equation (5) as follows,

$$r_t = \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \theta_i STMT_{t-5-i}) x_{t-i} + \varepsilon_{1,t},$$
$$x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \lambda_i STMT_{t-5-i}) x_{t-i} + \varepsilon_{2,t}, \quad (12)$$

TABLE 5 Vector autoregression estimation for Twitter sentiment three equations.

Panel A: Return equation	α_i Sum of lags	t-stat	β_i Sum of lags	t-stat	θ_i Sum of lags	t-stat	ζ_i Sum of lags	t-stat	η_i Sum of lags	t-stat	Adj-R^2
MessageVol	−0.087***	(−22.31)	0.445***	(67.31)	0.014***	(7.47)	0.031***	(10.44)	0.001	(0.31)	0.33
Bullishness	−0.086***	(−20.94)	0.448***	(66.47)	0.013***	(7.13)	0.006*	(1.94)	−0.003	(−1.18)	0.33
AgreementIndex	−0.086***	(−21.14)	0.449***	(68.35)	0.013***	(7.04)	0.003	(1.08)	−0.004	(−1.17)	0.33

Panel B: Trade equation	μ_i Sum of lags	t-stat	δ_i Sum of lags	t-stat	λ_i Sum of lags	t-stat	ι_i Sum of lags	t-stat	ν_i Sum of lags	t-stat	Adj-R^2
MessageVol	0.003	(0.33)	0.235***	(68.04)	−0.004*	(−1.81)	−0.027***	(−9.07)	−0.001	(−0.52)	0.05
Bullishness	0.001	(0.08)	0.235***	(68.78)	−0.003	(−1.20)	−0.001	(−0.33)	0.005	(1.60)	0.05
AgreementIndex	−0.001	(−0.10)	0.235***	(68.44)	−0.003	(−1.48)	−0.002	(−0.57)	0.002	(0.51)	0.05

Panel C: Limit order activity equation	Ω_i Sum of lags	t-stat	ξ_i Sum of lags	t-stat	π_i Sum of lags	t-stat	κ_i Sum of lags	t-stat	ψ_i Sum of lags	t-stat	Adj-R^2
MessageVol	−0.017***	(−4.40)	−0.009***	(−2.93)	0.717***	(130.27)	0.005**	(2.28)	0.024***	(9.06)	0.34
Bullishness	−0.018***	(−4.88)	−0.008**	(−2.47)	0.718***	(131.96)	−0.002	(−0.93)	0.010***	(3.54)	0.33
AgreementIndex	−0.018***	(−4.68)	−0.008**	(−2.53)	0.719***	(130.71)	0.003	(1.03)	0.012***	(4.24)	0.34

Note: This table reports the coefficient estimates of the VAR model for Twitter sentiment impacts. The estimation model is as below, the results for the return equation are presented in Panel A, the results for the trade equation are presented in Panel B, and the results for the limit order activity equation are presented in Panel C.

$$r_t = \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \theta_i STMT_{t-i}) x_{t-i} + \sum_{i=0}^{5} (\zeta_i + \eta_i STMT_{t-i}) QIT_{t-i} + \varepsilon_{1,t}$$
$$x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \lambda_i STMT_{t-i}) x_{t-i} + \sum_{i=1}^{5} (\iota_i + \nu_i STMT_{t-i}) QIT_{t-i} + \varepsilon_{2,t}$$
$$QIT_t = \sum_{i=1}^{5} \Omega_i r_{t-i} + \sum_{i=1}^{5} (\xi_i + \pi_i STMT_{t-i}) x_{t-i} + \sum_{i=1}^{5} (\kappa_i + \psi_i STMT_{t-i}) QIT_{t-i} + \varepsilon_{3,t}$$

where x_t is volume of signed order flow at the time interval t, r_t is trade return calculated from the natural logarithm of midquote at the t-th min, and $STMT_{t-i}$ is one of the Twitter sentiment measures: (i) *MessageVol*: the logarithm of the total number of tweets, (ii) *Bullishness*: a measure of investors' sentiment, or (iii) *AgreementIndex*: a measure of agreement among tweets. QIT_{t-i} is quote-intensity-to-trade ratio for the minute interval t, where quote intensity is defined as the summed number of changes in either price or depth at the best quotes on the limit order book. All variables are normalized. Newey-West corrected t-statistics are in parentheses. ***, **, and * indicate 1%, 5%, and 10% significance level, respectively.

TABLE 6 Vector autoregression estimation for non-contemporaneous Twitter sentiment.

Panel A: Return equation	α_i Sum of lags	t-stat	β_i Sum of lags	t-stat	θ_i Sum of lags	t-stat	Adj-R^2
MessageVol	−0.0748***	(−17.53)	0.4374***	(67.83)	0.0106***	(3.46)	0.33
Bullishness	−0.0744***	(−17.99)	0.4379***	(66.31)	0.0046*	(1.65)	0.33
AgreementIndex	−0.0744***	(−18.06)	0.4377***	(65.51)	0.0003	(0.10)	0.33
Panel B: Trade equation	μ_i **Sum of lags**	**t-stat**	δ_i **Sum of lags**	**t-stat**	λ_i **Sum of lags**	**t-stat**	**Adj-R^2**
MessageVol	−0.0014	(−0.15)	0.2456***	(68.06)	−0.0113***	(−3.37)	0.05
Bullishness	−0.0015	(−0.16)	0.2465***	(68.83)	0.0019	(0.56)	0.05
AgreementIndex	−0.0021	(−0.22)	0.2471***	(68.24)	−0.0022	(−0.69)	0.05

Note: This table reports the coefficient estimates of the VAR model for Twitter sentiment impacts. The estimation model is as below, the results of return equation are presented in Panel A, and the results of trade equation are presented in Panel B.

$$r_t = \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \theta_i STMT_{t-5-i}) x_{t-i} + \varepsilon_{1,t}$$

$$x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \lambda_i STMT_{t-5-i}) x_{t-i} + \varepsilon_{2,t}$$

where x_t is volume of signed order flow at the time interval t, r_t is trade return calculated from the natural logarithm of midquote at the t-th min, and $STMT_{t-i}$ is one of the following Twitter sentiment measures: (i) *MessageVol*: the logarithm of the total number of tweets, (ii) *Bullishness*: a measure of investors' sentiment, or (iii) *AgreementIndex*: a measure of agreement among tweets. All variables are normalized. Newey-West corrected t-statistics are in parentheses. ***, **, and * indicate 1%, 5%, and 10% significance level.

where the sentiment measures (*MessageVol*, *Bullishness*, or *AgreementIndex*) are lagged 5 min, $STMT_{t-5-i}$. That is, the sentiment is computed prior to the actual stock returns being realized. Consequently, it is not possible for future returns to influence past sentiment.

Table 6 reports the coefficients of the VAR system for the non-contemporaneous Twitter sentiment impacts on returns and trades in Panel A and B, respectively. We observe qualitatively similar results to our main findings in Table 2.

In Panel A, the summed coefficients of θ_i are positive and statistically significant at the 1% level for message volume and at the 10% level for bullishness. These coefficients are somehow weaker, suggesting that the effect of sentiment on return is stronger contemporaneously. More importantly, the instrumental variable analysis confirms that the impact of trades on SPY prices is affected by sentiment, that is, sentiment can influence intraday stock market returns through its impact on trades.

Panel B of Table 6 shows the non-contemporaneous effects of Twitter sentiment on the signed volume order flow. In line with Table 2, the summed coefficient of λ_i for *MessageVol* shows that when there are more tweets about SPY, the persistence in trades is less pronounced.

In summary, the findings of the non-contemporaneous effects of Twitter sentiment in Table 6 are similar to the contemporaneous effects in Table 2, and therefore, we conclude that Twitter sentiment affects stock market returns through its impact on trades.

5.2 | Using different thresholds when classifying tweets

Previously, we considered a tweet as positive (negative) if its tone score is above (below) zero. However, a tweet with a tone score of 0.1 (or −0.1), for instance, may not be distinguishable from a neutral tweet with a tone score of 0. These borderline tweets may affect the accuracy of our Twitter sentiment measure and weaken our results as the tweets we consider positive and negative may contain tweets that are, effectively, neutral-sounding. Thus, to ensure

TABLE 7 Vector autoregression estimation for Twitter sentiment ($POS > 0.3$ and $NEG < -0.3$).

	α_i		β_i		θ_i		
Panel A: Return equation	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Adj-R²*
MessageVol	−0.0736***	(−17.49)	0.4333***	(66.06)	0.0297***	(10.43)	0.33
Bullishness	−0.0748***	(−17.43)	0.4370***	(65.78)	0.0071**	(2.37)	0.33
AgreementIndex	−0.0747***	(−17.92)	0.4373***	(66.09)	0.0120***	(3.77)	0.33
	μ_i		δ_i		λ_i		
Panel B: Trade equation	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Adj-R²*
MessageVol	−0.0043	(−0.46)	0.2512***	(68.49)	−0.0249***	(−8.50)	0.05
Bullishness	−0.0040	(−0.45)	0.2496***	(69.88)	−0.0054*	(−1.82)	0.05
AgreementIndex	−0.0052	(−0.57)	0.2497***	(70.29)	−0.0095***	(−2.91)	0.05

Note: This table reports the coefficient estimates of the VAR model for Twitter sentiment impacts. The estimation model is as below, the results of return equation are presented in Panel A, and the results of trade equation are presented in Panel B.

$$r_t = \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \theta_i STMT_{t-i}) x_{t-i} + \varepsilon_{1,t}$$
$$x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \lambda_i STMT_{t-i}) x_{t-i} + \varepsilon_{2,t}$$

where x_t is volume of signed order flow at the time interval t, r_t is trade return calculated from the natural logarithm of midquote at the t-th min, and $STMT_{t-i}$ is one of the Twitter sentiment measures: (i) *MessageVol*: the logarithm of the total number of tweets, (ii) *Bullishness*: a measure of investors' sentiment, or (iii) *AgreementIndex*: a measure of agreement among tweets. All sentiment measures are based on positive and negative tweets classified using tone score greater than 0.3 and less than −0.3. All other scores are considered neutral tweets. All variables are normalized. Newey-West corrected t-statistics are in parentheses. ***, **, and * indicate 1%, 5%, and 10% significance level, respectively.

that our results are robust, we consider tweets with a tone score greater than 0.3 as positive and those with a score less than −0.3 as negative. The remaining tweets are considered neutral. We re-estimate Equation (5) based on the new sentiment classification.

Panel A of Table 7 shows that the sentiment coefficients, θ_i, are positive (0.0297 for *MessageVol*, 0.0071 for *Bullishness*, and 0.0120 for *AgreementIndex*) and statistically significant at the 5% level or better for the return equation. These coefficients are more positive and statistically significant compared to the same coefficients reported in Panel A of Table 2. Since we focus on tweets with more extreme values, the impact of sentiment is expected to be stronger, as indicated by the larger magnitudes of the coefficients.

Similarly, we observe that the sentiment coefficients λ_i are more negative and statistically significant in Panel B of Table 7 compared to those in Panel B of Table 2. For instance, the coefficient λ_i for *Bullishness* in Panel B of Table 7 is −0.0054 (significant at the 10% level), while the same coefficient in Panel B of Table 2 is −0.0016 but statistically insignificant. Furthermore, the coefficient λ_i for *AgreementIndex* in Panel B of Table 7 is −0.0095 (significant at the 1% level) versus −0.0024 in Table 2. Since λ_i represents the marginal impact of sentiment on trades, the lower $(\delta_i + \lambda_i)$ indicates less persistence in order flow if sentiment is more extreme. Combining the results in Table 7, we conclude that extreme tweets result in a larger price impact (Panel A) but weaker persistence of trades (Panel B), indicating less prevalence of trading continuation at intraday level.

5.3 | VAR with different lags

In all our VAR regressions, we employ five lags following Dufour and Engle (2000). To examine if our findings are sensitive to the choice of lags, we re-estimate Equation (5) using different sets of lags, for example, 3, 10, and 20 lags. We report the key results (interaction terms between sentiment and trades) in Table 8. They are qualitatively similar to

TABLE 8 Vector autoregression estimation for Twitter sentiment with different lags.

Panel A: Return equation	θ_i Sum of 3 lags	t-stat	θ_i Sum of 10 lags	t-stat	θ_i Sum of 20 lags	t-stat
MessageVol	0.0242***	(9.64)	0.0370***	(9.72)	0.0402***	(6.93)
Bullishness	0.0039	(1.58)	0.0090**	(2.28)	0.0166***	(2.91)
AgreementIndex	0.0020	(0.77)	0.0088**	(2.12)	0.0111*	(1.82)
Panel B: Trade equation	θ_i *Sum of 3 lags*	*t-stat*	θ_i *Sum of 10 lags*	*t-stat*	θ_i *Sum of 20 lags*	*t-stat*
MessageVol	−0.0180***	(−7.33)	−0.0394***	(−9.67)	−0.0587***	(−8.37)
Bullishness	−0.0022	(−0.92)	−0.0049	(−1.03)	−0.0108	(−1.55)
AgreementIndex	−0.0021	(−0.82)	−0.0035	(−0.75)	−0.0115*	(−1.66)

Note: This table reports the coefficient estimates of the VAR model for Twitter sentiment impacts using different lags. The estimation model is as below, the results of return equation are presented in Panel A, and the results of trade equation are presented in Panel B.

$$r_t = \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \theta_i STMT_{t-i}) x_{t-i} + \varepsilon_{1,t}$$
$$x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \lambda_i STMT_{t-i}) x_{t-i} + \varepsilon_{2,t}$$

where x_t is volume of signed order flow at the time interval t, r_t is trade return calculated from the natural logarithm of midquote at the t-th min, and $STMT_{t-i}$ is one of the Twitter sentiment measures: (i) *MessageVol*: the logarithm of the total number of tweets, (ii) *Bullishness*: a measure of investors' sentiment, or (iii) *AgreementIndex*: a measure of agreement among tweets. All variables are normalized. Newey-West corrected t-statistics are in parentheses. ***, **, and * indicate 1%, 5%, and 10% significance level, respectively.

those reported in Table 2, indicating that our results are robust to the choice of lags.

5.4 | Robustness with a different dictionary

Renault (2017) explains that the accuracy of sentiment categorization can influence the predictive power of sentiment measures. As our final robustness test, we use a different dictionary to classify tweets into positive, negative, and neutral groups. In particular, we use the Harvard IV-4 psychological dictionary to classify tweets.[15] We then re-construct our Twitter sentiment measures based on tweets classified using this dictionary and use them in our VAR model (Equation 5).

Table 9 reports the results using the new sentiment measures. Turning first to Panel A, we observe the mean reversion effect in intraday return (i.e., the sum coefficients of lagged returns) and the positive impact of order flow (i.e., the sum coefficients of the signed trading volume), consistent with the results in Table 2. For the interaction terms between Twitter sentiment and trades, we observe positive coefficients for message volume and agreement index (statistically significant at the 1% level). The sum of coefficients of bullishness remains positive, albeit insignificant. In Panel B of Table 9, for all three sentiment measures, the results are qualitatively similar to those reported in Panel B of Table 2.

It is important to note that when it comes to natural language processes, a context-related dictionary is more powerful than a general dictionary, as explained in Price et al. (2012). In line with this argument, there are two possible explanations why the sum of coefficients for bullishness is no longer significant after using the Harvard IV-4 psychological dictionary. First, WordNet employs a more extensive dictionary, encompassing 155,327 words, in stark contrast to

[15] Harvard IV-4 is a general-purpose dictionary developed by the Harvard University. According to its official documentary, the Harvard IV-4 is not particularly developed for natural language processing (NPL) objectives such as analyzing social media tweets as we do in this study. However, the Harvard IV-4 dictionary has been used for processing messages from official news media coverage (e.g., see Mangee, 2018; Price et al., 2012; Tetlock et al., 2008).

TABLE 9 Vector autoregression estimation for Twitter sentiment with a different dictionary.

	α_i		β_i		θ_i		
Panel A: Return equation	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Adj-R²*
MessageVol	−0.0736***	(−17.49)	0.4333***	(66.06)	0.0297***	(10.43)	0.33
Bullishness	−0.0751***	(−17.69)	0.4380***	(65.16)	0.0002	(0.05)	0.33
AgreementIndex	−0.0746***	(−18.06)	0.4381***	(64.90)	0.0082***	(2.61)	0.33
	μ_i		δ_i		λ_i		
Panel B: Trade equation	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Sum of lags*	*t-stat*	*Adj-R²*
MessageVol	−0.0043	(−0.46)	0.2512***	(68.49)	−0.0249***	(−8.50)	0.05
Bullishness	−0.0054	(−0.60)	0.2490***	(68.27)	−0.0050	(−1.56)	0.05
AgreementIndex	−0.0066	(−0.72)	0.2500***	(70.00)	−0.0038	(−1.07)	0.05

Note: This table reports the coefficient estimates of the VAR model for Twitter sentiment impacts where we use the Harvard IV-4 psychological dictionary to classify positive, negative and neutral tweets. The estimation model is as below, the results of return equation are presented in Panel A, and the results of trade equation are presented in Panel B.

$$r_t = \sum_{i=1}^{5} \alpha_i r_{t-i} + \sum_{i=0}^{5} (\beta_i + \theta_i STMT_{t-i}) x_{t-i} + \varepsilon_{1,t} x_t = \sum_{i=1}^{5} \mu_i r_{t-i} + \sum_{i=1}^{5} (\delta_i + \lambda_i STMT_{t-i}) x_{t-i} + \varepsilon_{2,t}$$

where x_t is volume of signed order flow at the time interval t, r_t is trade return calculated from the natural logarithm of midquote at the t-th min, and $STMT_{t-i}$ is one of the Twitter sentiment measures: (i) *MessageVol*: the logarithm of the total number of tweets, (ii) *Bullishness*: a measure of investors' sentiment, or (iii) *AgreementIndex*: a measure of agreement among tweets. All variables are normalized. Newey-West corrected t-statistics are in parentheses. ***, **, and * indicate 1%, 5%, and 10% significance level, respectively.

Harvard IV-4′s 4206 words. This allows WordNet to identify a broader spectrum of words within social media posts. Second, while Harvard IV-4 captures exact words, WordNet goes one step further by considering synonyms. Given these distinctions, we posit that WordNet excels in capturing retail investor sentiment, whereas Harvard IV-4 proves more adept in analyzing a formal style of writing. This may explain why *WordNet* is better suited for classifying texts in social media.[16] Appendix A offers a visual representation highlighting the distinctions between WordNet and Harvard IV-4.

6 | CONCLUSION

This study examines the price impact of tweets on stock markets. Using the SPDR S&P 500 Trust ETF (SPY) to proxy for the U.S. stock market, we explore the mechanism by which social media sentiment impacts SPY prices.

We find that Twitter sentiment enhances the impact of trades on intraday returns. Trades have a greater price impact when there are more tweets. Also, the impact of buyer- (seller-)initiated trades on stock prices is stronger when the market sentiment is bullish (bearish). The impact of Twitter sentiment on prices causes a lingering mispricing effect that is not fully assimilated at the intraday level. Rather, this mispricing takes several days to correct. This price impact is not subsumed by market volatility and liquidity shocks. We also test if Twitter sentiment affects prices through limit order activity but do not find that is the case. Our results highlight the importance of social media sentiment in stock market price movements at the intraday level.

[16] According to the General Inquirer, for creating a helpful category for content analysis, the extensive electronic lexical database such as Wordnet can be considerable for improvement. See, for instance, https://inquirer.sites.fas.harvard.edu/homecat.htm.

ACKNOWLEDGMENTS

We thank Vitali Alekseev, Henk Berkman, Bart Frijns, Nhut (Nick) Nguyen Hoang, Dinh Phan, Andranik Tumasjan, Nuttawat Visaltanachoti, Peiming Wang, the editors Michael Pagano and Tina Yang, two anonymous referees, and participants at the 25th New Zealand Finance Colloquium, the 11th Financial Markets and Corporate Governance Conference and Auckland University of Technology finance seminar for their valuable comments and suggestions.

Open access publishing facilitated by Auckland University of Technology, as part of the Wiley - Auckland University of Technology agreement via the Council of Australian University Librarians.

ORCID

Adrian Fernandez-Perez https://orcid.org/0000-0002-9433-7194
Ivan Indriawan https://orcid.org/0000-0002-6558-5551

REFERENCES

Agrawal, S., Azar, P. D., Lo, A. W., & Singh, T. (2018). Momentum, mean-reversion, and social media: Evidence from StockTwits and Twitter. *Journal of Portfolio Management*, *44*(7), 85–95.

Akhtar, S., Faff, R., Oliver, B., & Subrahmanyam, A. (2011). The power of bad: The negativity bias in Australian consumer sentiment announcements on stock returns. *Journal of Banking and Finance*, *35*(5), 1239–1249.

Akhtar, S., Faff, R., Oliver, B., & Subrahmanyam, A. (2012). Stock salience and the asymmetric market effect of consumer sentiment news. *Journal of Banking and Finance*, *36*(12), 3289–3301.

AlMousa, M., Benlamri, R., & Khoury, R. (2021). Exploiting non-taxonomic relations for measuring semantic similarity and relatedness in WordNet. *Knowledge-Based Systems*, *212*, 106565.

Al-Nasseri, A., Ali, F. M., & Tucker, A. (2021). Investor sentiment and the dispersion of stock returns: Evidence based on the social network of investors. *International Review of Financial Analysis*, *78*, 101910.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56.

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance*, *59*(3), 1259–1294.

Azar, P. D., & Lo, A. W. (2016). The wisdom of Twitter crowds: Predicting stock market reactions to FOMC meetings via Twitter feeds. *Journal of Portfolio Management*, *42*(5), 123–134.

Baker, M., & Stein, J. C. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets*, *7*(3), 271–299.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance*, *61*(4), 1645–1680.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, *21*(2), 129–152.

Bali, T. G., Peng, L., Shen, Y., & Tang, Y. (2014). Liquidity shocks and stock market reactions. *Review of Financial Studies*, *27*(5), 1434–1485.

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, *21*(2), 785–818.

Behrendt, S., & Schmidt, A. (2018). The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility. *Journal of Banking and Finance*, *96*, 355–367.

Bhala, R. V., & Abirami, S. (2014). Trends in word sense disambiguation. *Artificial Intelligence Review*, *42*(2), 159–171.

Birru, J. (2018). Day of the week and the cross-section of returns. *Journal of Financial Economics*, *130*(1), 182–214.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1–8.

Broadstock, D. C., & Zhang, D. (2019). Social-media and intraday stock returns: The pricing power of sentiment. *Finance Research Letters*, *30*, 116–123.

Brogaard, J., Hendershott, T., & Riordan, R. (2019). Price discovery without trading: Evidence from limit orders. *Journal of Finance*, *74*(4), 1621–1658.

Brown, P., Walsh, D., & Yuen, A. (1997). The interaction between order imbalance and stock price. *Pacific-Basin Finance Journal*, *5*(5), 539–557.

Chen, H., Noronha, G., & Singal, V. (2004). The price response to S&P 500 index additions and deletions: Evidence of asymmetry and a new explanation. *Journal of Finance*, *59*(4), 1901–1930.

Chordia, T., Roll, R., & Subrahmanyam, A. (2001). Market liquidity and trading activity. *Journal of Finance*, *56*(2), 501–530.

Chordia, T., Roll, R., & Subrahmanyam, A. (2005). Evidence on the speed of convergence to market efficiency. *Journal of Financial Economics*, *76*(2), 271–292.

Conrad, F. G., Gagnon-Bartsch, J. A., Ferg, R. A., Schober, M. F., Pasek, J., & Hou, E. (2021). Social media as an alternative to surveys of opinions about the economy. *Social Science Computer Review*, *39*(4), 489–508.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, *98*(4), 703–738.

Dougal, C., Engelberg, J., Garcia, D., & Parsons, C. A. (2012). Journalists and the stock market. *Review of Financial Studies*, *25*(3), 639–679.

Dufour, A., & Engle, R. F. (2000). Time and the price impact of a trade. *Journal of Finance*, *55*(6), 2467–2498.

Dumas, B., Kurshev, A., & Uppal, R. (2009). Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. *Journal of Finance*, *64*(2), 579–629.

Edmans, A., Fernandez-Perez, A., Garel, A., & Indriawan, I. (2022). Music sentiment and stock returns around the world. *Journal of Financial Economics*, *145*, 234–254.

Edmans, A., Garcia, D., & Norli, Ø. (2007). Sports sentiment and stock returns. *Journal of Finance*, *62*(4), 1967–1998.

Gan, B., Alexeev, V., Bird, R., & Yeung, D. (2020). Sensitivity to sentiment: News vs social media. *International Review of Financial Analysis*, *67*, 101390.

Gao, Z., Ren, H., & Zhang, B. (2020). Googling investor sentiment around the world. *Journal of Financial and Quantitative Analysis*, *55*(2), 549–580.

Goettler, R. L., Parlour, C. A., & Rajan, U. (2009). Informed traders and limit order markets. *Journal of Financial Economics*, *93*(1), 67–87.

Goetzmann, W. N., Kim, D., Kumar, A., & Wang, Q. (2015). Weather-induced mood, institutional investors, and stock returns. *Review of Financial Studies*, *28*, 73–111.

Gu, C., & Kurov, A. (2020). Informational role of social media: Evidence from Twitter sentiment. *Journal of Banking & Finance*, *121*, 105969.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance*, *46*(1), 179–207.

Hirshleifer, D., Jiang, D., & DiGiovanni, Y. M. (2020). Mood beta and seasonalities in stock returns. *Journal of Financial Economics*, *137*(1), 272–295.

Hirshleifer, D., & Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *Journal of Finance*, *58*, 1009–1032.

Hoffmann, P. (2014). A dynamic limit order market with fast and slow traders. *Journal of Financial Economics*, *113*(1), 156–169.

Huang, H. G., Tsai, W. C., Weng, P. S., & Wu, M. H. (2021). Volatility of order imbalance of institutional traders and expected asset returns: Evidence from Taiwan. *Journal of Financial Markets*, *52*, 100546.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, *95*(1), 161–182.

Jordà, Ò. (2023). Local projections for applied economics. *Annual Review of Economics*, *15*, 607–631.

Kamstra, M. J., Kramer, L. A., & Levi, D. (2000). Losing sleep at the market: The daylight-savings anomaly. *American Economic Review*, *90*, 1005–1011.

Kamstra, M. J., Kramer, L. A., & Levi, M. D. (2003). Winter blues: A SAD stock market cycle. *American Economic Review*, *93*(1), 324–343.

Karagozoglu, A. K., & Fabozzi, F. J. (2017). Volatility wisdom of social media crowds. *Journal of Portfolio Management*, *43*(2), 136–151.

Kim, S. H., & Kim, D. (2014). Investor sentiment from internet message postings and the predictability of stock returns. *Journal of Economic Behavior and Organization*, *107*, 708–729.

Kocoń, J., & Maziarz, M. (2021). Mapping WordNet onto human brain connectome in emotion processing and semantic similarity recognition. *Information Processing and Management*, *58*(3), 102530.

Kordonis, J., Symeonidis, S., & Arampatzis, A. (2016). Stock price forecasting via sentiment analysis on Twitter. *Proceedings of the 20th pan-hellenic conference on informatics*, *36*, 1–6.

Kurov, A. (2008). Investor sentiment, trading behavior and informational efficiency in index futures markets. *Financial Review*, *43*(1), 107–127.

Lee, C. M., & Ready, M. J. (1991). Inferring trade direction from intraday data. *Journal of Finance*, *46*(2), 733–746.

Lee, R. S. (2023). Semantic analysis. In *Natural language processing* (pp. 115–147). Springer.

Liu, S. (2015). Investor sentiment and stock market liquidity. *Journal of Behavioral Finance*, *16*(1), 51–67.

Mangee, N. (2018). Stock returns and the tone of marketplace information: Does context matter? *Journal of Behavioral Finance*, *19*(4), 396–406.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *The Journal of Finance*, *32*(4), 1151–1168.

Moseki, K. K., & Rao, K. M. (2018). Empirical measures of symmetry of market sentiments. *Cogent Economics & Finance*, *6*(1), 1430113.

Navigli, R. (2009). Word sense disambiguation: A survey. *ACM Computing Surveys (CSUR)*, *41*(2), 1–69.

Price, S. M., Doran, J. S., Peterson, D. R., & Bliss, B. A. (2012). Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking and Finance*, *36*(4), 992–1011.

Rakowski, D., Shirley, S. E., & Stark, J. R. (2021). Twitter activity, investor attention, and the diffusion of information. *Financial Management*, *50*(1), 3–46.

Renault, T. (2017). Intraday online investor sentiment and return patterns in the U.S. stock market. *Journal of Banking and Finance*, *84*, 25–40.

Roşu, I. (2019). Fast and slow informed trading. *Journal of Financial Markets*, *43*, 1–30.

Schnaubelt, M., Fischer, T. G., & Krauss, C. (2020). Separating the signal from the noise–financial machine learning for Twitter. *Journal of Economic Dynamics and Control*, *114*, 103895.

Schober, M. F., Pasek, J., Guggenheim, L., Lampe, C., & Conrad, F. G. (2016). Social media analyses for social measurement. *Public Opinion Quarterly*, *80*(1), 180–211.

SEC. (2013). *SEC says social media OK for company announcements if investors are alerted*. SEC. https://www.sec.gov/news/press-release/2013-2013-51htm

Shu, H. C., & Chang, J. H. (2015). Investor sentiment and financial market volatility. *Journal of Behavioral Finance*, *16*(3), 206–219.

Siganos, A., Vagenas-Nanos, E., & Verwijmeren, P. (2014). Facebook's daily sentiment and international stock markets. *Journal of Economic Behavior and Organization*, *107*, 730–743.

Sprenger, T. O., Sandner, P. G., Tumasjan, A., & Welpe, I. M. (2014). News or noise? Using Twitter to identify and understand company-specific news flow. *Journal of Business Finance and Accounting*, *41*(7-8), 791–830.

Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and trades: The information content of stock microblogs. *European Financial Management*, *20*(5), 926–957.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, *104*(2), 288–302.

Sun, L., Najand, M., & Shen, J. (2016). Stock return predictability and investor sentiment: A high-frequency perspective. *Journal of Banking and Finance*, *73*, 147–164.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, *62*(3), 1139–1168.

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance*, *63*(3), 1437–1467.

Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: Survey-based indicators vs. Google trends. *Journal of Forecasting*, *30*(6), 565–578.

How to cite this article: Yang, N., Fernandez-Perez, A., & Indriawan, I. (2025). The Price Impact of Tweets: A High-Frequency Study. *Financial Review*, *60*, 147–171. https://doi.org/10.1111/fire.12406

APPENDIX A

Diagram of the various words related to the keyword "book" in WordNet

This diagram (Lee, 2023) demonstrates the network when detecting the word "book" in WordNet. The words are interconnected by conceptual lexical and semantic relations, meaning not only word forms such as strings of letters are considered but also senses of words.